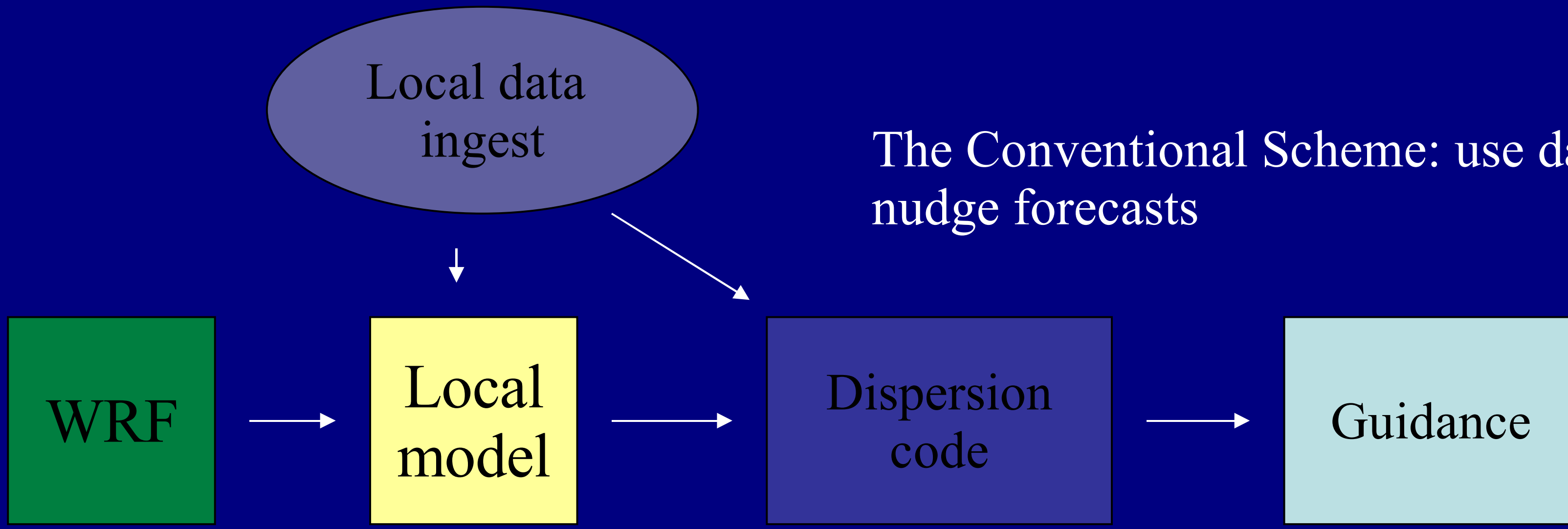
Urban Model Output Statistics

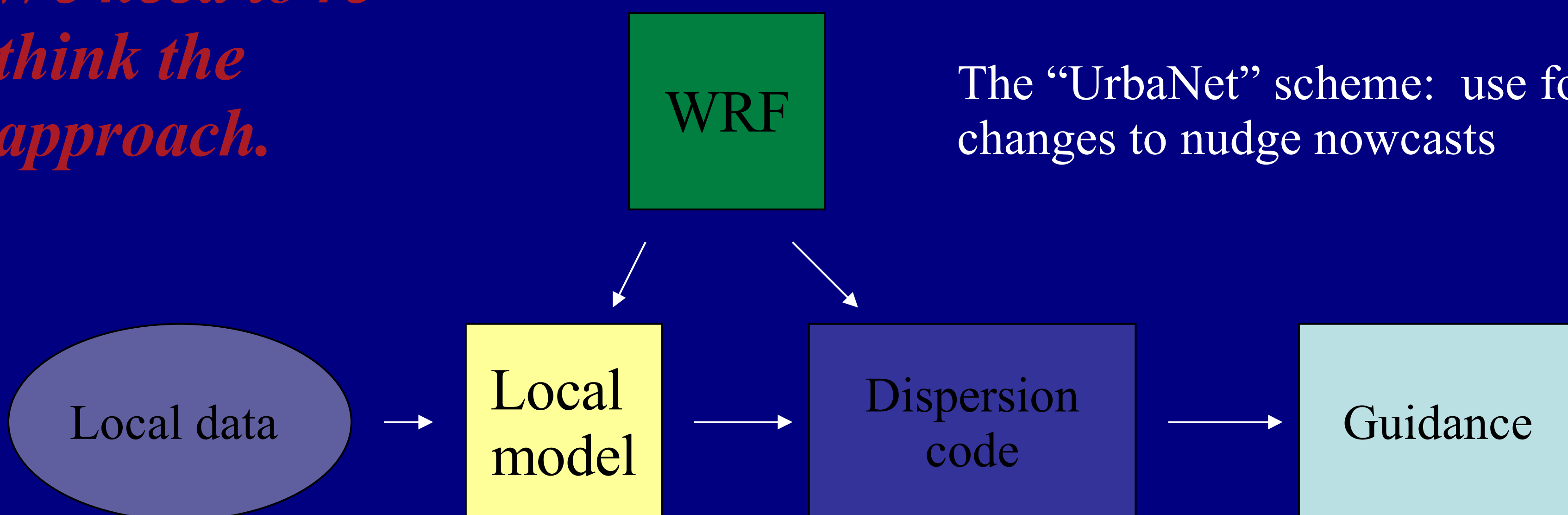
Richard M. Eckman
NOAA/ARL Field Research Division
Idaho Falls



The Conventional Scheme: use data to nudge forecasts

This works well on the average. For events, it often fails.

We need to re-think the approach.



The “UrbaNet” scheme: use forecast changes to nudge nowcasts

Probabilistic Forecast Uncertainty

F = Quantity we want to predict
M = Model estimate of quantity
O_a = Observations used to initialize model

$$p(F|O_a) = \int p(F|M)\, p(M|O_a)\, dM$$

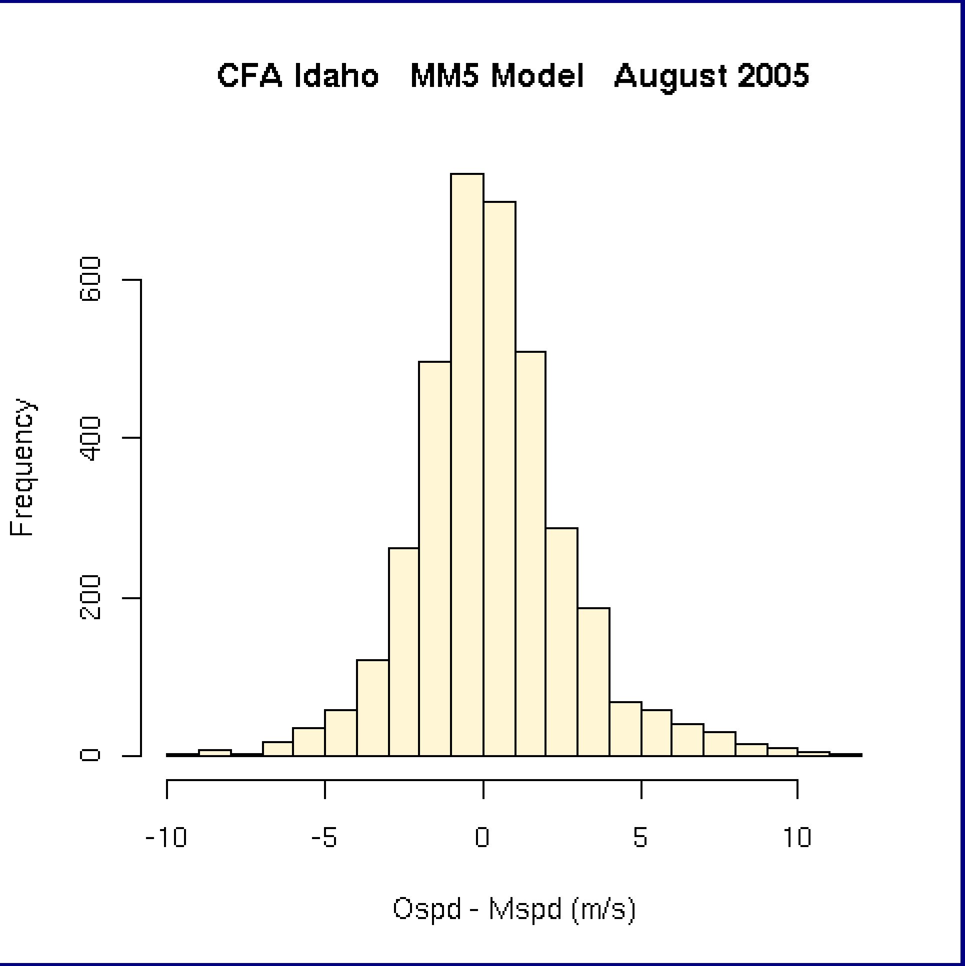
Total uncertainty

Model physics errors

Model input uncertainty (ensemble fcst)

Mean value theorem: $p(F|O_a) = p(F|\hat{M})$

Observed Speeds vs MM5 Forecast



Effect of Recent Observations

O_r = Recent observations not included in model initialization

$$p(F|O_a, O_r) = \int p(F|M, O_r)\, p(M|O_a)\, dM$$

Additional conditional information (O_r) hopefully decreases uncertainty

Conventional Model Output Statistics (MOS)

- Normal linear model

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots \beta_k x_{ik} + \epsilon_i$$

or

$$\boldsymbol{y} = \boldsymbol{X} \boldsymbol{\beta} + \boldsymbol{\epsilon}$$

- X variables are combination of model estimates and recent observations
- Coefficients $\boldsymbol{\beta}$ are different for each station
- Direct attempt to model $p(F|M, O_r)$

Example MOS Equation

12 Z wind speed at St. Louis
Glahn and Lowry (1972)

$$S_f^{12} = 1.576 + 0.239\,S_m^{12} + 0.175\,S_o^{07} - 0.040\,V_m^{12} + 0.027\,U_m^{12}$$

Combines both 12 Z model output and recent observations at 7 Z

Urban MOS

$$y = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

- y = urban dispersion parameters (skimming flow, turbulence levels)
- X = model outputs (RUC) + local observations (e.g., private network data)
- May use blend of many local observations.
- Utility of private network data demonstrated by improvements in regression

Predictive Distribution

- Predictive distribution for y is multivariate t
- Provides both point estimate and prediction uncertainty
- Standard MOS usage focuses only on point estimate

Advantages

- Conceptually simple
- Linked to private network data
- NOAA operations familiar with MOS approach
- Uses statistics to account for physics not resolved in model
- Predictive distribution gives an estimate of uncertainty in forecast

Disadvantages

- Might need a lot of different regression equations
- Difficult finding reasonable number of explanatory variables
- Might need hierarchical structure to account for complications (e.g., $\boldsymbol{\beta}$ a function of hyperparameters)
- Purists may criticize lack of physics in regression